2024 3rd International Conference for Advancement in Technology (ICONAT)
Goa, India. Sep 13-14, 2024

Secure Virtual Desktop Infrastructure Solution Using Homomorphic Encryption and Machine Learning Models

Senuwan W.M.D
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
IT21118036@my.sliit.lk

Nimantha Dissanayake
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
IT21097560@my.sliit.lk

Malithi Disanayaka
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
IT21010026@my.sliit.lk

Shashini Hewage
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
IT21035562@my.sliit.lk

Kavinga Yapa Abeywardena
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
kavinga.y@sliit.lk

Deemantha Siriwardhana
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
deemantha.s@sliit.lk

***Abstract*— In the recent past, Virtual Desktop Infrastructure (VDI) technology has experienced rapid growth. Although many enterprise companies have implemented VDI systems, previous research has highlighted several critical issues, including data breaches and session hijacking. Security measures identified for such threats were suboptimal, especially for the firms with fewer resources. In this paper, authors combat existing threats to VDI systems with different technologies such as Homomorphic Encryption (HE) for the safe exchange of location data, and the Machine Learning (ML) model is used for real-time log analysis, thereby making it easy to detect some level of anomaly in the complex VDI environments. According to the result, this integrated posture introduces a new and positive dimension in VDI strengthening the security and privacy of data, while protecting against various threats. Additionally, the paper also addresses setting a new and more secure base for any virtualization technologies all while securely handling consumer data.**

Keywords—Virtual Desktop Infrastructure (VDI), Homomorphic Encryption (HE), Location Based Service (LBS), Anomality detection, Logs, Machine Learning (ML), Data Center

I. INTRODUCTION

VDI is a technology that can facilitate the creation, management, and delivery of virtualize desktops from a centralized sever. This architectural approach allows users to access their desktop environments remotely, providing flexibility and efficiency in virtualized desktop management. By hosting desktop operating systems such as Windows, on a server within a Data Center. Users can connect to these virtual desktops which are pre-assigned by administrators via various devices such as laptops, mobile phones, or tablets, over the internet.

In the recent past, VDI technology has grown at an extreme pace and has altered the ways through which it can be implemented in the digital landscape of an organization. VDI solutions are replacing conventional remote desktop solutions in many sectors and companies, for having the advantages, Centralized Management allows administrators to manage multiple virtual desktops simultaneously, and apply updates and patches efficiently. While an organization can easily scale its desktop infrastructure up or down based on demand, often leveraging cloud services to reduce hardware costs can be achieved. As VDI technology assures almost zero latency, users can access their virtual desktops from anywhere, facilitating remote work more efficiently and effectively.

By 2027, the global market share for VDI solutions is expected to reach USD 4.5 billion [1]. This suggested that the use of VDI solutions is growing and it is necessary to secure the Virtualized Infrastructure. However, industries focused on how well VDI solutions performed and whether or not they could be effectively and successfully implemented, all while neglecting fundamental security concerns such as Data Leakage and session hijacking, which will be addressed in the paper.

This research contributes to this understanding by proposing novel approaches based on Homomorphic Encryption (HE) for the exchange of location data. HE is a technique that allows computation to be performed on encrypted data (Ciphertext) without needing to decrypt (Plaintext). The outcome of computing encrypted data will be the same as the outcome of computing plaintext values using the same computational.

However, Logs are records of a sequence of activities and events that transpire in systems, applications, and even on the network level, hence critical to the health and overall security and performance of the IT environment. Since organizations are using complex and dynamic systems and relying on them more and more for log analysis, the authors have introduced a novel approach for detecting anomalies in real-time with a trained ML model, thus improving the current state of the VDI.

This study therefore defines an important research area by concentrating on the integration of VDI technology and improved security protocols, while addressing the gap in the current perspective about the best practices related to the protection of privacy and the ability to detect intrusions.

979-8-3503-5417-1/24/$31.00 ©2024 IEEE

II. Background & Literature Review

VDI term coined by VMWare back in 2006. This was the first provider of VDI solutions and later started to provide virtualized apps. Later Citrix and Microsoft adopted the power of providing virtualized apps and later became VDI vendors. As with the competition, the rapid development and advancement of the VDI technology began. However, there is particularly very limited research being conducted on securing VDI solutions.

Tong et.al [2] in 2015, proposed a Virtual Local Area Network (VLAN) -based VDI architecture to address concerns in data security and enterprise infrastructure protection. Their research focuses on how VLANs can segregate and isolate network traffic in a VDI environment, thereby reducing the attack surface and enhancing the security of both user data and the overall network infrastructure. By such VLAN implementation in a VDI Setup, Tong et.al aims to ensure that even if one segment of the network had compromised, the breach would not necessarily spread to other branches of the infrastructure, thus providing a more secure environment for enterprise data.

Yoo [5] in 2012, conducted a case study at a university hospital to examine the practical challenges associated with implementing a VDI system. This study highlighted several key issues, including the architectural design of the VDI system, the integration of wireless network infrastructure, and the need for strong user authentication mechanisms. They found that the implementation of VDI in a healthcare environment posed unique challenges due to the sensitivity of patient data and the requirement for high levels of security and reliability. They mentioned that, deploying robust authentication methods to verify user identities and using wireless protocols to protect data transmitted over the network.

Thus, the above security solution has been limited to the backend processes and architectural solutions. There is a limited solution to protect user privacy and detect intrusion in the VDI environments. However, to fulfill this gap the proposed solutions were analyzed. The concept of “Privacy Homomorphisms” [13] is not new. It was first introduced by Rivest's work, as a method for key exchange within the system. Their research concluded with open questions about the practicality of implementation or circuity reconfiguration needed for successful operations.

Another significant development in this area was Pailler’s work [14] on the composite residual class problem. Unlike Rivest’s approach [13], Pascal Pailler proposed a novel mathematical problem that enabled additive homomorphism on ciphertexts, in contrast to the integer factorization problem. Which only supported the multiplication of two ciphertexts. Pailler’s cryptosystem was later classified as a “Partial Homomorphic Scheme”.

Boneh later developed a “Somewhat Homomorphic” scheme that allowed unlimited additive operations and a single multiplication operation on ciphertexts. Their work [15] resulted in a constant-size ciphertext that could support these operations. However, in 2009, Craig Gentry introduced the concept of Fully Homomorphic Encryption (FHE) [8], which allows computations to be performed on encrypted data without decryption [16]. This paved the breakthrough the way for secure computation of sensitive data in various applications. Even though, the practical challenges do exist with limitations of implementing FHE schemes [7]. In a discussion in 2011, it was highlighted that a need for further research to improve the efficiency and practicality of HE for real-world application.

However, these studies demonstrate the potential use of HE in enhancing the data security and privacy of a user. However, the method was deemed usable, and secure. In which his work [8], a noise factor was added during the process. Additionally, the scheme allowed all operations to be performed within the ciphertext. These schemes were utilized by their respective authors to address the general issues of retaining privacy while computing. Herein assess the fundamental use of mentioned schemes to preserve the privacy of users, and do computation on the data for analysis. Since there are proven methods to analyze and process encrypted data [9,10]. As a result, the technology can be used for safe data computation in unmanaged or outsourced contexts without jeopardizing users' privacy this paper will address the performance and scalability challenges associated with implementing HE in VDI deployments.

Due to the rising complexity and frequency of cyber-attacks, there is a greater demand for real-time log analysis for anomaly detection utilizing machine learning. In a work [11], they have proposed a technique based on Knowledge Discover Database (KDD) datasets that outperform existing machine learning techniques and could be utilized to develop intrusion detection strategies. As well as using supervised learning techniques such as Muti-layer Perceptron, Support Vendor Machine (SVM) and rule-based systems outperformed others in detecting abnormalities [12].

III. Research Gap

Open-source VDI solutions such as Apache Guacamole [6] have been a popular player for a long period. With the simplicity of deployment, organizations tend to use it. However, enterprises with resources deploy such open-source VDI backends, as they equip secure monitoring systems and logging tools. Even though, some enterprise tools do not consider the privacy of consumers. In one instance, it was noted that a world-recognized tool known as CyberArk enables administrators to view the location of the users.

Location-based services (LBS) services collect raw data about consumers' geo-location. These data are stored in a database for future analysis. At the moment, this data is being used to be processed and do the follow-up analysis on the consumer trends to detect anomalies and to recommend services on the geo-location. Implementing encryption for such databases can be implemented. Moreover, several types of research have been conducted on breaking this cryptography with the emergence of the era of quantum computers. However, the encryption technique was controlled only when the data was at rest, after which data was decrypted at the stage of data processing for analysis. To mitigate such data leakage, the paper proposes a HE based mechanism, which can ensure that consumer data will be encrypted at all times. With such advanced integration of HE, the feasibility of storing such data in cloud environments securely is an additional advantage.

Even though, there are various tools and techniques to detect intrusion detection, the majority of VDI systems do not utilize the Zero Trust Security Model, which is considered more and more as a standard in the field of cybersecurity. The Zero Trust model starts from the fact that threats may come

from outside and inside, it focuses on access planning and constant monitoring. If these principles are not integrated, then VDI systems can still be insecure from inside threats in addition to outside threats. Also, the response and remedial measures in current systems incorporated in recent technology are not proactive enough with regards to responding to threats as they are found; they can only respond to the threat in a mechanistic manner and may take a long time to initiate proper measures to counteract or remedy the situation.

TABLE I. COMPARISON OF CURRENT SYSTEMS AND THE PROPOSED SYSTEM

Feature	Current Systems	Proposed System
Adaptive Learning Algorithms	✗	✓
Predictive Analysis Capabilities	✗	✓
Automated Response and Detection	✗	✓

IV. METHODOLOGY

VDI solutions are sensitive working environments, and securing such infrastructure can be a hard task to achieve, as VDI solutions to run effectively the concept of zero-latency must be maintained. For this, the proposed system will be focused on user location activity and user behavior analysis. This comprehensive approach includes several interconnected components, the following sub-sections provide an in-depth analysis of the techniques and a summary of the specific functionality of each component proposed.

A. Tracking Users's Location While Preserving Privacy

The architecture of the proposed LBS consists of the following components

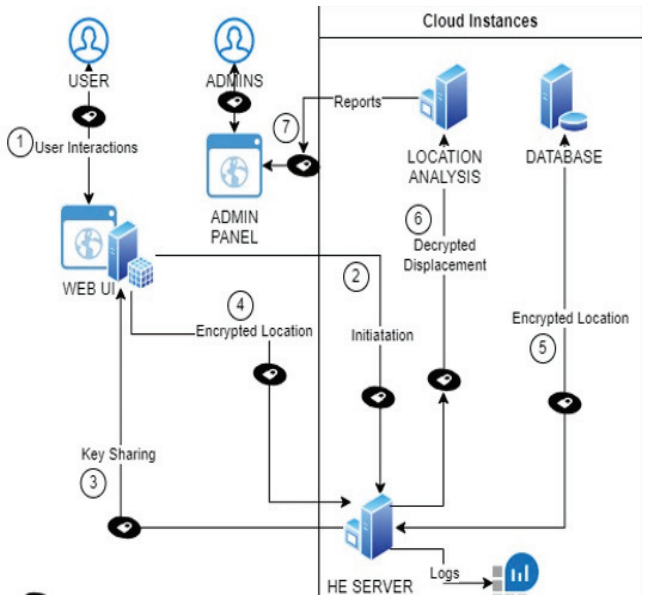

Fig. 1. System Diagram

"USER" is the endpoint where users interact with the interface, services will request the longitude and latitude of the user, and the user will be provided with encrypted location data.

There is a process that needs to be followed before data encryption, to encrypt data using HE, data needs to be encoded. There are several methods to perform encoding. However, the use of HE and CKKS (Cheon-Kim-Kim-Song) schemes limits these options. When considering SEAL Libraries, it was developed to support Single instruction/Multiple Data (SIMD). This utilized parallel processing which enables more efficient operations. Native encoder which was embedded in the library supports this technology. This can be utilized for recurring updates of user location flows. It is suitable for the proposed system, as it can be scaled up depending on the consecutive requests that were received by the server to perform, HE functions.

After which cipher context of the algorithm must be initiated. The cipher context includes values, parameters, and arguments that will be used to initialize the HE processes successfully. Before selecting these parameters, encryption requirements must be thoroughly analyzed. This includes the purpose, data that are being used, plaintext data types, and storage requirements. It is a vital step since the cipher context is initialized based on these requirements.

Values are considered integers or chars. The remaining float values will be rounded to base 20 to get the integer. Since the location values are in the float format, the CKKS scheme is used as the encryption scheme as the value is not fixed and float values are used. However, there was considerable storage overhead associated with the CKKS scheme compared to BFV (Brakerski-Fan-Vercauteren), even though the CKKS scheme is the only scheme that can operate on the real number domain. It possesses many capabilities compared to BFV. The polynomial modulo and its related coefficient modulo are adjusted according to performance and accuracy.

Secure communication channels such as HTTPS will be used for transmitting encrypted location data to the HE server. HE server at the backend that received encrypted location data will store the data in the HE database. The database can be utilized to process data for computation on it. Any algorithms that are optimized or built with the CKKS scheme are compatible with this. The current scheme supports addition and multiplication operations. This is facilitated by the evaluator function in the library. Since these values are encrypted, the leakage of such a database can result in no harm for most of the part as homomorphic encryption is a quantum resistance encryption. Following up, unlike encrypting a whole database to a different cipher context will create a severe processing load on the system.

The Computation module will perform computations and will be presented to the analytics module to analyze displacement and anomalies. The technique used for displacement computation is based on the Euclidean Distance Formula [17] (Refer to Equation 1). Here "D" is the Euclidean distance while "x" represents the Longitude and "y" represents the latitude. Moreover, x_1, y_1 will be the initial geo-location details, and x_2, y_2 will be the second geo-location details.

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Equation 1 - Euclidean Distance Formula

Therefore, considering every last decimal bit was very important when considering obtaining longitude and latitude values, as we are processing this information to detect anomaly detection of users based on their location.

However, computing on the set of the same data locations as the consumers will be at rest can generate noise, which is also known as the “Noise Factor”. This is addressed, with the use of the “Relinearization” methodology. Moreover, there were methodologies such as “Bootstrapping”, which was compared to the current method allowing a more reduced and refreshed budget which can be less in size compared to Relinearization operation. However, it hasn’t been implemented in SEAL, which is for future developmental improvements, as it is in the SEAL’s roadmap for CKKS. That being highlighted, implementing bootstrapping can cause a significant performance hit, often a large magnitude in resource usage. Considering the requirement for the process requires accuracy, this is highly important and must be taken into consideration for future developments.

B. Advanced Real-time log analysis and anomaly detection

In the proposed system, real-time log analysis for anomaly detection refers to the practice of continuously examining and reviewing log data as it happens to uncover unexpected patterns or behaviors that may reveal a security threat, system failure, or illegal activity. The business applies deterministic advice for finding out symptoms by logs that are simple to discover using normal keywords of key phrases such as “error,” “failure,” or “warning.” An anomaly is a term that can be quickly distinguished, and the solution comprises data accumulation, and preprocessing, as well as machine learning models for identifying anomalies that are supervised, unsupervised, and semi-supervised learning models.

ML types have different uses and optimal performance under certain situations, and this is the significance of the article. These systems include the requirement of an appropriate event log dataset. Since there is no matching dataset for this system to match the relevant requirements, a matching dataset was created and used for this system., and various system events were tracked to detect anomalies. It captures essential details like event types, timestamps, user information, and risk levels, which are crucial for identifying patterns of normal and suspicious activities. The dataset is labeled with different risk categories such as "Low," "Medium," and "High." These labels are based on predefined rules or criteria that assess the severity and likelihood of the event being associated with a security threat, making it suitable for supervised learning models that classify events as either normal or abnormal. By including user and system details, along with the outcomes of each event, the dataset enables a comprehensive analysis that helps train models to spot unusual behaviors. This is essential for building an effective anomaly detection system that can quickly identify and address potential security threats in VDI environments. Several suspicious activities that mainly affect Windows Remote machines were used for this. It sends alerts to suitable teams concerning anomalies signed by the ML models that distinguish normal and anomalous data.

This system diagram gives information about an anomaly detection system, which utilizes advanced machine learning and filters the event logs stored in the system. The advanced machine learning model is trained with the event logs stored in the system. After analyzing the event data stored in this

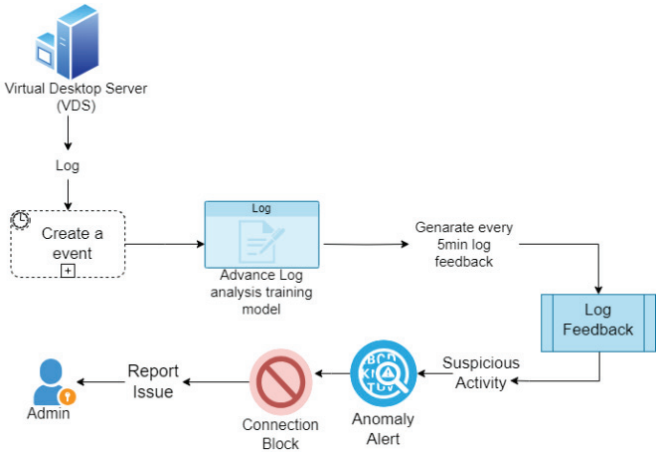

Fig. 2. Realtime Log analysis diagram

way, event feedback is given in 5-minute intervals. Through the Event Feedback given in this way, suspicious activities in the system will be identified. Then the identified suspicious activities are automatically terminated by the system. In this way, every activity that is terminated will notify the admin through an alert. While, the system handles threats quickly, and the admin is always informed and ready to take further action if needed.

The sensitive information in the system should be protected from most authorized users. The use case that comes under the Anomaly detection system is the Immediate Threat Detection and Response, Reduced False Positives, Continuous Monitoring, Proactive Threat Hunting, Operational Efficiency and Performance Monitoring, Data Privacy and Security. However, in the process of developing the real-time log analysis and anomaly detection system for VDI, the authors decided to use it. Because the system is specifically developed to operate on the Windows platform, I propose to use the NET as the framework. .NET is a framework that was developed by Microsoft and as such it is well suited for developing applications on Windows.

Since a large portion of the VDI environment is based on Windows, it is not advisable to use. The suitability of NET is to make sure that the system fully integrates with the established structures. This choice helps the system to parse logs; identify problematic patterns and act on them in case of potential security risks in real time along with leveraging the strengths of the Windows architecture. In short, .NET was chosen to ensure that the system operates efficiently and optimally in a Windows-based VDI infrastructure.

The system has trained the dataset using the Random Forest model because it gave the highest accuracy. We also attempted such models as Logistic Regression, LDA, and the Naive Bayes, however, the performance of such models was subpar. We used Random Forest as our model for this project since it achieved the highest accuracy among all the compared algorithms. In the management of security risks in events, this research has grouped them into four security risk levels to ensure quick identification of risks and subsequent response to the same. List of High-Risk events Privilege escalation, unusual login attempts, Event log clearing, and Critical also belong to high risks. The Medium-Risk event New or unknown process, unusual network activity, or Error. Low-risk events are File and directory changes, unexpected system changes, and Warnings. Normal events, which are unspectacular although crucial, are informational and must be

continually observed. This helps in focusing on the actions to take and or prevent further destruction of system security. In our risk management, there is something referred to as 'Medium Risk' activities; these are not ideals for high risk but should not be taken lightly. For example, a medium risk may be categorized by the emergence of new or unidentified processes, irregular networking, or errors that are beyond the threshold of the system. Nevertheless, the timing of such activities is extremely important. If medium-risk events happen during non-business hours such as late evening or night-time, then they need to be considered as critical risks. This is because processes that are not regular, deviate from normal network traffic, or the errors that occur during odd hours are probably not part of normal business processes and therefore possibly indicative of a security breach. Hence, it is crucial to pay a lot of attention to the activities within such a period since they may signify either an ongoing or a likely security threat.

V. Results & Discussions

The system was able to deploy state of the art cryptographic approaches, new means of authentications, and incorporate several machine learning algorithms to enhance the security of Virtual Desktop Infrastructure (VDI) settings.

While the Random Forest model for real-time log analysis and detecting anomalies tracked a very high overall accuracy of 99%. The classification report which followed yielded near perfect precision, recall and F1-scores for the different classes with classes 1 and 3 being perfectly scored with an F1 of 100%. The macro and weighted averages supported and strengthened the belong model, weights averages make nearly perfect at 99%. (Refer Table II)

TABLE II. RF Classification Report

	Precision	Recall	F1-score	support
0	0.94	1.00	0.97	36333
1	1.00	1.00	1.00	361634
2	0.41	0.03	0.06	2262
3	1.00	1.00	1.00	21703
Accuracy			0.99	421932
Macro Avg	0.84	0.76	0.76	421932
Weight Avg	0.99	0.99	0.99	421932

The research also presented an efficient HE-based method to solve security challenges in Location-Based Services (LBS) while ensuring user privacy gets preserved but still offering efficient services. (Refer Figures 3, 4 & 5)

```
curl -X 'POST' \
  'https://localhost:7067/Location/AddNewLocation' \
  -H 'accept: */*' \
  -H 'Content-Type: application/json' \
  -d '{
  "latitude": 6.883705665893483,
  "longitude": 79.97682720232041,
  "userName": "test3"
}'
```

Fig. 3. Initial Position

```
curl -X 'POST' \
  'https://localhost:7067/Location/AddNewLocation' \
  -H 'accept: */*' \
  -H 'Content-Type: application/json' \
  -d '{
  "latitude": 6.838801639470187,
  "longitude": 79.98074214868484,
  "userName": "test3"
}'
```

Fig. 4. Second position

Fig. 5. Calculated Displacement

The above figures illustrate an example use case, between the Athurugiriya Interchange and the Hiripitya Road Interchange, which is almost a linear road stretch of 5.1km on Google Maps. As we are calculating displacement, the above value depicted in Figure 5 is 98.2% accurate relative to the distance shown by Google Maps.

These studies demonstrated how integrating state-of-art machine learning algorithm such as the Random Forest classifier with other state-of-art cryptographic as well as authentication methods. This integration also improved cloud-based VDI security and paved the way for other advancements in the field of cybersecurity and privacy.

VI. Conclusion

This research investigates the problem associated with the increasing use of Virtual Desktop Infrastructure (VDI) systems and poor security. Here, we utilized HE, and real-time analysis to achieve secure data sharing without compromising the user's privacy. As the new generation threats continue to emerge, our approach uses HE-based location analysis of logs in real-time combined with a simple authentication process in addition to a fast-real-time log analyzer that lays VDI solutions to strengthen against hackers.

In this case, the innovation is in the broad approach which not only complies with standard regulations but also dampens the effects of cyber threats. Further work will be done in adding new components and interlinking them to form a system of systems where overall performance is improved through aggregation of data at the component level and better anomaly detection. Furthermore, by adopting homomorphic encryption for the processing of users' data, consent management will be more trustworthy.

Thus, our solution does have some limitations, most significantly for the currently supported geolocation data fields that are encrypted, which consist only of latitude and longitude. This could be taken further into future developments where extra geolocation parameters are included but this always poses the problem of latency. Nevertheless, the study presents a solid ground upon which subsequent investigations of the security of VDI solutions can be built.

REFERENCES

[1] Food Community Circle, "Virtual Desktop Infrastructure (VDI) Market Share," LinkedIn Pulse, https://www.linkedin.com/pulse/virtual-desktop-infrastructure-vdi-market-share/. [Accessed: Dec 11, 2023].

[2] Y. J. Tong, W. Q. Yan, and J. Yu, "Analysis of a Secure Virtual Desktop Infrastructure System," International Journal of Digital Crime and Forensics, vol. 7, pp. 69-84, 2015. doi: 10.4018/IJDCF.2015010105.

[3] "Securing Your Workforce: The Role of Virtual Desktop Infrastructure in Cybersecurity," Apporto, 2023. [Online]. Available: https://www.apporto.com/securing-your-workforce-the-role-of-virtual-desktop-infrastructure-in-cybersecurity.

[4] "Virtual Desktop Infrastructure (VDI) Security Risks," OPSWAT, 2024. [Online]. Available: https://www.opswat.com/blog/virtual-desktop-infrastructure-vdi-security-risks.

[5] "Identifying and Analyzing Security Threats to Virtualized Cloud," IEEE Xplore, 2013. [Online]. Available: https://ieeexplore.ieee.org/document/6488089.

[6] Apache Software Foundation, "Apache Guacamole: Access your computers from anywhere," [Online]. Available: https://guacamole.apache.org.

[7] DataPort. [Online]. Available: https://ieee- dataport.org/sites/default/files/analysis/27/IEEE%20Citation%20Guidelines.pdf. [Accessed: May. 16, 2024].

[8] H.-I. Kim, H.-J. Kim, and J.-W. Chang, "A secure kNN query processing algorithm using homomorphic encryption on outsourced database," Data and Knowledge Engineering, vol. 123, p. 101602, Sep. 2019, doi: 10.1016/j.datak.2017.07.005. Available: https://doi.org/10.1016/j.datak.2017.07.005

[9] Carmen Lee, "Privacy-preserving proof-of-location using homomorphic encryption", Uppsala Universitiet, June.2020. [Accessed Jan 15, 2024].

[10] Wikipedia contributors, "Euclidean distance," Wikipedia, The Free Encyclopedia,Available:https://en.wikipedia.org/wiki/Euclidean_distance. [Accessed: March 9, 2024].

[11] Jabez, J., & Muthukumar, B. (2015). Intrusion Detection System (IDS): Anomaly detection using Outlier Detection Approach. Procedia Computer Science, 48, 338-346.

[12] Wahba, M., Farouk, A., & Abozaid, S. (2015). Intrusion detection system using feature selection based on ensemble learning algorithms. Procedia Computer Science, 65, 36-45.

[13] D. Crosbie, "An Exploration of Zero-Knowledge Proofs and zk-SNARKs," 2019

[14] D. Boneh, E.-J. Goh, and K. Nissim, "Evaluating 2-DNF formulas on ciphertexts," in Lecture Notes in Computer Science, 2005, pp. 325–341

[15] C. Gentry, "Fully homomorphic encryption using ideal lattices," STOC '09: Symposium on Theory of Computing, May 2009, doi: 10.1145/1536414.1536440.Available: https://doi.org/10.1145/1536414.1536440

[16] L. Pulido-Gaytan, A. Tchernykh, J. Cortes-Mendoza, M. Babenko and ´ G. Radchenko," A Survey on Privacy-Preserving Machine Learning with Fully Homomorphic Encryption," in Latin America High-Performance Computing, Cuenca, Ecuador, 2020

[17] C. Helzner, "Euclidean Distance | Calculation, Formula & Examples," Study.com. [Online]. Available: https://study.com/academy/lesson/euclidean-distance-calculation-formula-examples.html.